Solution to Exercise 9.5

⋆9.5 Consider the linear regression model with serially correlated errors,

$$y_t = \beta_1 + \beta_2 x_t + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \tag{9.117}$$

where the ε_t are IID, and the autoregressive parameter ρ is assumed either to be known or to be estimated consistently. The explanatory variable x_t is assumed to be contemporaneously correlated with ε_t (see Section 8.4 for the definition of contemporaneous correlation).

Recall from Chapter 7 that the covariance matrix $\boldsymbol{\Omega}$ of the vector $\boldsymbol{u}$ with typical element u_t is given by (7.32), and that $\boldsymbol{\Omega}^{-1}$ can be expressed as $\boldsymbol{\Psi}\boldsymbol{\Psi}^\top$, where $\boldsymbol{\Psi}$ is defined in (7.60). Express the model (9.117) in the form (9.20), without taking account of the first observation.

Let Ω_t be the information set for observation t with $\mathrm{E}(\varepsilon_t \,|\, \Omega_t) = 0$. Suppose that there exists a matrix $\boldsymbol{Z}$ of instrumental variables, with $\boldsymbol{Z}_t \in \Omega_t$, such that the explanatory vector $\boldsymbol{x}$ with typical element x_t is related to the instruments by the equation

$$\boldsymbol{x} = \boldsymbol{Z}\boldsymbol{\pi} + \boldsymbol{v}, \tag{9.118}$$

where $\mathrm{E}(v_t \,|\, \Omega_t) = 0$. Derive the explicit form of the expression $(\boldsymbol{\Psi}^\top\bar{\boldsymbol{X}})_t$ defined implicitly by equation (9.24) for the model (9.117). Find a matrix $\boldsymbol{W}$ of instruments that satisfy the predeterminedness condition in the form (9.30) and that lead to asymptotically efficient estimates of the parameters β_1 and β_2 computed on the basis of the theoretical moment conditions (9.31) with your choice of $\boldsymbol{W}$.

As usual, we denote by $\boldsymbol{y}$ the n–vector with typical element y_t. Using expression (7.60) for $\boldsymbol{\Psi}$, we see that $(\boldsymbol{\Psi}^\top\boldsymbol{y})_t = y_t - \rho y_{t-1}$ for $t = 2, \ldots, n$. Thus, ignoring the first observation, we can write the model (9.117) in the form of equation (9.20) as follows:

$$y_t - \rho y_{t-1} = (1-\rho)\beta_1 + \beta_2(x_t - \rho x_{t-1}) + \varepsilon_t, \tag{S9.06}$$

since $(\boldsymbol{\Psi}^\top\boldsymbol{u})_t = u_t - \rho u_{t-1} = \varepsilon_t$ by the specification (9.117).

According to equation (9.24), $\mathrm{E}\big((\boldsymbol{\Psi}^\top\boldsymbol{X})_t \,|\, \Omega_t\big) = (\boldsymbol{\Psi}^\top\bar{\boldsymbol{X}})_t$. In what follows, we ignore the constant in (S9.06), since it is clear that the instruments $\boldsymbol{W}$ must in any case include a constant. Thus we restrict our attention to $\mathrm{E}\big((\boldsymbol{\Psi}^\top\boldsymbol{x})_t \,|\, \Omega_t\big)$, which, by (9.118), is

$$\begin{aligned}
\mathrm{E}\big((\boldsymbol{\Psi}^\top\boldsymbol{x})_t \,|\, \Omega_t\big) &= \mathrm{E}(x_t - \rho x_{t-1} \,|\, \Omega_t) \\
&= \mathrm{E}(\boldsymbol{Z}_t\boldsymbol{\pi} + v_t - \rho x_{t-1} \,|\, \Omega_t) \\
&= \boldsymbol{Z}_t\boldsymbol{\pi} - \rho x_{t-1}.
\end{aligned}$$

The last equality uses the facts that $\mathrm{E}(v_t \,|\, \Omega_t) = 0$ by assumption and that $x_{t-1} \in \Omega_t$, since x_{t-1} is predetermined for observation t. It follows that $(\boldsymbol{\Psi}^\top \bar{\boldsymbol{x}})_t = \boldsymbol{Z}_t \boldsymbol{\pi} - \rho x_{t-1}$, and so $\boldsymbol{\Psi}^\top \bar{\boldsymbol{x}} = \boldsymbol{Z}\boldsymbol{\pi} - \rho \boldsymbol{x}_{-1}$, where $(\boldsymbol{x}_{-1})_t = x_{t-1}$.

We now see that $\boldsymbol{\Psi}^\top \bar{\boldsymbol{x}} \in \mathcal{S}(\boldsymbol{Z}, \boldsymbol{x}_{-1})$. Since $\boldsymbol{Z}_t \in \Omega_t$ and $\boldsymbol{x}_{-1} \in \Omega_t$, the instruments $\boldsymbol{W}$, except for the constant, may be taken to be such that $\boldsymbol{\Psi}^\top \boldsymbol{W} = [\boldsymbol{Z} \;\; \boldsymbol{x}_{-1}]$. Formally, $\boldsymbol{W} = (\boldsymbol{\Psi}^\top)^{-1} [\boldsymbol{Z} \;\; \boldsymbol{x}_{-1}]$.

The moment conditions (9.31) are such that it is not necessary to use the inverse of $\boldsymbol{\Psi}^\top$ explicitly, since $\boldsymbol{W}$ enters only through the transpose of $\boldsymbol{\Psi}^\top \boldsymbol{W}$. Using (S9.06) in order to get an explicit expression for $\boldsymbol{\Psi}^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$ in (9.31), we find that the efficient estimates are obtained by using the theoretical moment conditions that $\varepsilon_t = y_t - \rho y_{t-1} - (1-\rho)\beta_1 - \beta_2 x_t + \rho \beta_2 x_{t-1}$ should be orthogonal to a constant term, $\boldsymbol{Z}_t$, and x_{t-1}. Thus if $\boldsymbol{A} \equiv [\boldsymbol{\iota} \;\; \boldsymbol{Z} \;\; \boldsymbol{x}_{-1}]$, the empirical moments are

$$\boldsymbol{A}^\top \big(\boldsymbol{y} - \rho \boldsymbol{y}_{-1} - (1-\rho)\beta_1 - \beta_2 \boldsymbol{x} + \rho \beta_2 \boldsymbol{x}_{-1} \big),$$

and the optimal weighting matrix is proportional to $(\boldsymbol{A}^\top \boldsymbol{A})^{-1}$.